A Web-based semantic tagging and activity recognition system for species’ accelerometry data

Lianli Gao, Hamish A. Campbell, Owen R. Bidder, Jane Hunter

School of ITEE, The University of Queensland, Brisbane, QLD 4072, Australia
School of Biological Sciences, The University of Queensland, Brisbane, QLD 4072, Australia
Biological Sciences, College of Science, Swansea University, Singleton Park, Swansea SA2 8PP, UK

Abstract

Increasingly, animal biologists are taking advantage of low cost micro-sensor technology, by deploying accelerometers to monitor the behavior and movement of a broad range of species. The result is an avalanche of complex tri-axial accelerometer data streams that capture observations and measurements of a wide range of animal body motion and posture parameters. Analysis of these parameters enables the identification of specific animal behaviors—however the analysis process is immature with much of the activity identification steps undertaken manually and subjectively. Consequently, there is an urgent need for the development of new tools to streamline the management, analysis, indexing, querying and visualization of such data. In this paper, we present a Semantic Annotation and Activity Recognition (SAAR) system which supports storing, visualizing, annotating and automatic recognition of tri-axial accelerometer data streams by integrating semantic annotation and visualization services with Support Vector Machine (SVM) techniques. The interactive Web interface enables biologists to visualize and correlate 3D accelerometer data streams with associated video streams. It also enables domain experts to accurately annotate or tag segments of tri-axial accelerometer data streams, with standardized terms from an activity ontology. These annotated data streams can then be used to dynamically train a hierarchical SVM activity classification model, which can then be applied to new accelerometer data streams to automatically recognize specific activities. This paper describes the design, implementation and functional details of the SAAR system and the results of the evaluation experiments that assess the performance, usability and efficiency of the system. The evaluation results indicate that the SAAR system enables ecologists with little knowledge of machine learning techniques to collaboratively build classification models with high levels of accuracy, sensitivity, precision and specificity.

© 2012 Elsevier B.V. All rights reserved.
accelerometer data (Halsey et al., 2009), empirical calculation via
threshold values (Lagarde et al., 2008) or mathematical calculations
over Microsoft Excel spreadsheets (Shepard et al., 2008). Some recent
studies have applied machine learning techniques to recognize activities
in humans (Zhenyu and Lianwen, 2009) and animals (Martiskainena et
al., 2009; Nathan et al., 2012). However, this research focuses on the
optimization of feature extraction algorithms rather than the provi-
sion of an integrated Web-based system that combines visualization,
tagging and automatic recognition tools with a repository for storing
the indexed tri-axial accelerometer data streams.

In particular, the interpretation of wild animal accelerometer data is
a difficult problem due to: the volume and complexity of the data
streams; the variability of activity and behavioral patterns across an-
imals (due to age, environment, season); the lack of visualization
and analysis tools; the inability to share data and knowledge between ex-
erts; the lack of verifiable reference data e.g., using observational
video; and the inaccessibility of automatic recognition services. More-
over, the steep learning curve associated with building and applying
machine learning or pattern recognition techniques to accelerometer
data limits the accessibility of these approaches to a relatively small
group of experts.

The SAAR (Semantic Annotation and Activity Recognition) system
was designed to provide solutions to the issues outlined above and to
achieve the following objectives and user requirements:

• To provide a repository on the Web where researchers monitoring
animal behavior, can upload and share their datasets—and also
search, retrieve and compare datasets from the same or different
species;
• To provide interactive graphical visualization services that enable
scientists to quickly and easily view and explore tri-axial acceler-
ometer data streams and temporally align simultaneously re-
corded video (where available) that can be used to verify behavior/
activities;
• To provide a platform by which ecologists can interactively record,
share and re-use domain expert knowledge on animal movements
within tri-axial accelerometer data streams in an interoperable,
re-usable manner;
• To provide a set of Web services that can be used to analyze, tag and
visualize 3D accelerometer datasets and synchronized video using
terms from controlled vocabularies (pre-defined ontologies);
• To enable ecologists to build their own automatic activity recognition
models by training classifiers using features extracted from pre-
annotated training sets;
• To assess the quality of results generated by Support Vector Machine
(SVM)-based activity recognition classifiers that have been trained
using manually annotated data for a variety of species (human, dog,
badger);
• To determine whether an activity recognition classifier trained using
data from one species (e.g., a domestic dog) can be usefully applied to
other species (e.g., a badger) or wild species (e.g., a dingo), of similar
size and gait.
• To enable the sharing, re-use and refinement of activity recognition
classifiers developed for specific species, between scientists.

In the remainder of this paper, we describe the SAAR system in
more detail. Section 2 provides detailed information about the re-
search methodology, the underlying data/metadata models and the
automatic machine learning approach we have adopted. Section 3 de-
scribes the implementation details and provides information about
the system infrastructure, functionality and screen shots of the user
interface. Section 4 describes the evaluation process. Section 5 analy-
ses the evaluation results including the performance of the hierarchi-
cal classifier and the usability of the system. Finally, we discuss our
contributions, system limitations, as well as the future plans and
conclusions.

2. Methodology

2.1. Case study and data collection

The challenge for many ecologists is to understand the movement
and behavior of animals “in the wild”. Researchers are currently using
accelerometers to measure the activity levels and movement of many
wild animals (including crocodiles (Campbell et al., 2010), bears
(Gervasi et al., 2006) and badgers (Shepard et al., 2008)) to assist
with their management and conservation. In Australia, researchers
are investigating the behavior and movement of wild dogs and
dinges in order to develop appropriate management strategies
(DERM, 2006; Mitchell and Balogh, 2007). The difficulty with analyz-
ing tri-axial accelerometer data from wild animals is that there is lit-
tle or no observational data or video that provides the evidence for
training an automatic activity recognition model. One of the secondary
aims of our work described here is to develop a model for domestic
quadruped mammals (i.e., domestic dogs) using the associated video
as verification and to see if this model can be used to accurately recog-
nize activities of other similar-sized quadruped mammals (e.g., bad-
gers) or similar species in the wild (e.g., dingo), for which there is no
corresponding observational video.

We collected data from eight voluntary students and staff from the
University of Queensland (four males and four females). They are all
healthy with no physical problems and their ages range from 25 to
38. This human data collection was divided into two stages: training
data collection and test data collection. The G6A was attached to the
shoulder of each volunteer with X axis pointing backwards, Y axis
pointing left and Z axis pointing upward. At the training data collec-
tion stage, each volunteer was asked to do 3 minute walking, 3 min-
ute running, 1 minute sit-ups, 3 minute standing, 3 minute sitting and
3 minute lying. At the test data collection stage, each volunteer was
asked to arbitrarily perform these same six activities over a
15 minute period. During the entire data collection phase, a camera
simultaneously recorded video which provides the ground truth for
the evaluation phase.

Next, we collected data from domestic dogs. The same accelerom-
ter device (G6A) was attached to the back of each dog’s neck via its
collar. Six dogs of different breeds and ages were observed (4 year old
Border Collie, 15 kg weight, 52 cm height; 1 year old dachshund,
8.9 kg weight, 20 cm height; 8 year old Cocker Spaniel, 14 kg weight,
35 cm height; 5 year old German Short-Haired Pointer, 25.8 kg
weight, 63 cm height; 10 years old Staffordshire Terrier–Labrador
cross, 21 kg weight, 55 cm height; 5 years old Cavalier King Charles
Spaniel, 7.5 kg weight, 30 cm height). During the training data collec-
tion stage, each dog was led by its owner to perform 2 minute walk-
ing, 2 minute running, 2 minute standing, 2 minute sitting and
2 minute lying. In addition the King Charles Spaniel spent 1 minute
foraging/digging and 1 minute climbing (front paws raised to reach a
treat, while the owner walked backwards). During the test data collec-
tion stage, each dog was led by its owner to randomly perform the
activities above over a period of 10 min. During the entire data collec-
tion phase, a camera simultaneously recorded video which provides
the ground truth for the evaluation phase.

In addition, Eurasian badger data was collected from studies un-
taken at West Hatch RSPCA Centre, Somerset, UK. During these
studies, five Eurasian badgers were equipped with tri-axial acceler-
ometers that were attached to a leather collar fastened round the
badgers’ necks with X axis pointing backwards, Y axis pointing left and Z axis pointing upward (Shepard et al., 2008). Camera traps were also set up to verify some activities, although large periods of activity were outside the camera’s field of view. Where no verification by video was possible, manual annotations were made based on prior knowledge and the principals set out by Shepard et al. (2008). Six activities were annotated: walking, running, climbing, foraging, standing and lying.

2.2. Process

Our approach to this research can be sub-divided into the eight stages described below:

1. A Web interface was developed that enables datasets (tri-axial accelerometer data in CSV format and corresponding videos in Ogg Vorbis OGV format) to be uploaded to the system’s server and described using simple metadata including: Creator, dateCaptured, Species, AnimalID, Location, Coordinates, Description.
2. Users can search, browse, retrieve and open specific datasets and visualize both the tri-axial accelerometer data (and associated video if available) through a graphical user interface that comprises two panels (Plot and Video) – juxtaposed one above the other – that display both the tri-axial movement data streams and the video stream. Simple alignment tools enable users to precisely synchronize the data and video streams.
3. An ontology-based annotation service enables domain experts to tag tri-axial accelerometer data streams manually via the combined Plot and Video user interface—using predefined ontologies that capture the terms describing activities of interest to the researcher and of relevance to the animal being studied e.g. running, walking, standing, sitting, and lying. Separate ontologies can be developed for different terrestrial, marine and avian species and the most appropriate ontology selected at run-time.
4. The manually attached tags (and pointers to relevant file segments/time stamps) are stored on an annotation server in RDF format. Through the annotation interface, users can share their tags with other users, search and retrieve specific tags/annotations and associated accelerometer data segments e.g. give me all segments in which animal with ID “abcd’” is “running”.
5. A user then specifies the set of tagged data streams which are to be used as the training data. The system retrieves and aggregates all of the data corresponding to each tag/label and extracts a set of application-dependent features that represent that tag/label. The application-dependent features and representative labels are then used to interactively train a hierarchical SVM classifier that recognizes both “active” and “inactive” states as well as more specific sub-class activities.
6. When new tri-axial accelerometer data streams are uploaded, the corresponding application-dependent features are extracted and then input into the trained SVM classifier which automatically tags/annotates the new data streams. The classification results are stored in RDF on the annotation server and displayed via the Web visualization interface for biologists to verify or correct.
7. Finally, statistical analysis tools are also provided that calculate the statistics for each activity for a single animal or a set of animals (including average, minimum, maximum time of occurrence, cumulative time of occurrence in the whole period, total number of occurrence and standard deviation of the duration time). These results are presented as a pie chart on the Web interface.
8. To evaluate the system we assess the performance of different SVM classifiers by comparing the automatically tagged data streams against ground truth data (captured via video or hand tagged data streams). We also assess the system’s usability and efficiency by collecting and analyzing users’ feedback and performance metrics.

2.3. Activity recognition using support vector machines

SVMs (Support Vector Machines) are well established as a successful modeling and prediction tool for both pattern classification and regression tasks. They are linear classifiers based on statistical learning theory and the idea of the maximum margin hyper-plane. In previous species activity identification studies (Martiskainen et al., 2009; Zhen-Yu and Lian-Wen, 2008, 2009), SVMs demonstrate relatively good performance when applied to the classification of tri-axial accelerometer data streams from humans and cows. For SAAR, we decided to use the LIBSVM library (Chang and Lin, 2011) because it is open source, written in Java and is simple to download and use. More specifically, we decided to use the C-SVC (C-support Vector Classification) algorithm (Boser et al., 1992) from the LIBSVM library because it is the simplest SVM approach.

Our activity recognition service is designed to perform on two hierarchical levels: high level recognition and low level recognition. The high level recognition service identifies active and inactive activities, while the low level recognition recognizes specific activities which are sub-classes of the active and the inactive activity classes (for example, walking, running, feeding, sleeping, lying, etc.). In order to use the C-SVC algorithms to automatically recognize tri-axial accelerometer data stream patterns, application-dependent features have to be extracted.

In this study, features were extracted using a window size of 3 s with an overlap of 1 s (2 sampling points for 1 Hz sampling rate) between consecutive windows. There are three reasons for selecting this window length and overlap. Firstly, feature extraction on sliding windows with 50% overlap (2/4 samples overlap) has been demonstrated to achieve accurate results in previous research efforts (Bao and Intille, 2004; Li et al., 2010; Ravi et al., 2005; Yang et al., 2008). Secondly, it has been shown that a window of 2 s can capture activities (Li et al., 2010), hence a window of 3 s with 1 second overlap will be sufficient to capture activities. Thirdly, the most efficient algorithm for calculating the Fast Fourier Transform (FFT) usually operates with a time window length that is a power of two.

At the high level recognition, we extract the following features including standard deviation vector, signal magnitude area vector and waveform length vector. They are expressed respectively as follows:

- **Standard deviation (SD):** The standard deviation measures how spread out the signal is within x-axis, y-axis and z-axis respectively.

\[ SD = \sqrt{\frac{1}{N-N^{-1}} \sum_{i=1}^{N} \left( x_i - \frac{1}{N} \sum_{k=1}^{N} x_k \right)^2} \]

Where \( x_i \) and \( x_k \) are the ith and the kth accelerometer values on the x-axis, y-axis and z-axis, and N is the window size.

- **Signal magnitude area (SMA):** The signal magnitude area is found to be a suitable measurement of the degree of movement intensity that can distinguish between active and inactive activities using tri-axial accelerometer data (Khan et al., 2008).

\[ SMA = \frac{1}{N} \left( \sum_{i=1}^{N} |x_i| + \sum_{i=1}^{N} |y_i| + \sum_{i=1}^{N} |z_i| \right) \]

- **Waveform Length (WL):** The WL is the cumulative length of the waveform amplitude, frequency and duration all within a signal window.
In other words, it measures the total amount variance of signal vibration through three dimensions.

\[ WL = \frac{1}{N-1} \sum_{i=1}^{N-1} |x_i - \bar{x}| + \sum_{i=1}^{N-1} |y_i - \bar{y}| + \sum_{i=1}^{N-1} |z_i - \bar{z}| \]

For the low level recognition, we extract spatial-domain features (standard deviation, signal magnitude area, waveform length). But in addition we extract frequency-domain features and an inheritance parameter.

The discrete Fourier transform (DFT), a transform for Fourier analysis of finite-domain discrete-time signals, is widely employed in signal processing to produce frequency information contained in a sampled signal (Smith, 1999). A fast Fourier transform (FFT) is an efficient algorithm to compute DFT and it produces exactly same results of DFT (Brigham, 1988). Given a set of real or complex numbers \(x_0, ..., x_{N-1}\), the DFT transforms them into the sequence of \(N\) complex numbers \(X_0, ..., X_{N-1}\). Those complex numbers represents the magnitude and phrase information about the transformed sequence. In this study, we take the power of magnitude of the complex FFT output as the component of the frequency-domain features.

Fig. 1 illustrates how to use FFT transform to compute frequency-domain features for the low level activity recognition. The Inheritance parameter (IP) measures whether a subclass was originally inherited from a parent class. To compute the IP value, high level activity recognition is employed to recognize two classes: active activity and the inactive activity. The value of IP is 1 if the classification result belongs to the active activity class, and −1 if it belongs to the inactive activity class.

3. Implementation

3.1. System architecture

Fig. 2 shows the high level architectural components of the SAAR system\(^1\) which combines: Web 2.0 technologies (Java, JavaScript, and JSON) to maximize accessibility and collaboration; with Semantic Web technologies (RDF, SPARQL, OWL ontologies) to maximize knowledge capture, re-use and interoperability through standardized vocabularies; and Support Vector Machine (SVM) to provide the machine-learning tools for automated recognition of activities.

A Web-based Plot-Video visualization interface combines AJAX, Flot (a plotting library),\(^2\) HTML 5 Video Player library (Video.js)\(^3\) with JavaScript to enable users to interactively visualize both tri-axial accelerometer data alongside simultaneously recorded videos in an interactive plot visualization pane and a video player, respectively.

Using the Plot-Video visualization interface, users can invoke the semantic annotation service by selecting a segment of tri-axial accelerometer data alongside simultaneously recorded videos in an interactive plot visualization pane and a video player, respectively.

The activity recognition is implemented using the LIBSVM Java library. At the training stage, users interactively search and retrieve specific segments/annotations via the following search terms: species, creator, animal ID and activityTag. The SPARQL query language is used to query the annotation server and automatically transform the retrieved annotations into a set of application-dependent features with representative labels based on users’ activity recognition level selection. After the specific hierarchical SVM classification model is built for all of the activity tags, new tri-axial accelerometer data are input to the trained SVM classifier to automatically tag the input data. The predicted results are displayed in the timeline visualization pane, where experts can check or correct them. An advanced statistical analysis of animal activity information is conducted on the predicted results by using simple statistical algorithms and the results are displayed in a 3D pie chart.

3.2. User interface

The SAAR user interface, accessible via a Firefox or Chrome Web browser, enables users to interactively:

- Zoom in or zoom out the timeline visualization interface to precisely attach an activity tag to a segment of tri-axial accelerometer data streams (motion along the X, Y and Z axes);
- Synchronize the video player with the timeline visualization so users can attach a tag/annotation to either a segment of tri-axial accelerometer data stream or the video and it is attached to both segments;
- Delete, edit or correct annotations;
- Search and retrieve annotations based on annotation content and metadata. For example: give me all annotations created by a user “Juana” between the “2012-03-01 00:00:00” and “2012-03-02 00:00:00”;
- Dynamically train an SVM activity classifier using annotated data streams and then apply this trained classification model to newly generated accelerometer data streams to automatically tag activities;
- Statistically analyze the tags on a data stream to calculate relative times spent by a particular animal or species on each activity.

Fig. 3 illustrates the SAAR Plot-Video visualization interface and the annotation interface. The top left of the interface shows the Plot
Fig. 2. High level architectural view of the SAAR system.

Fig. 3. Screenshot of SAAR Plot-Video visualization interface and the annotation interface.
interface and the tri-axial accelerometer data stream (for a domestic dog). The X-axis data is yellow, the Y-axis data is blue and, and Z-axis data is red. Users are able to zoom in and zoom out to observe the data streams in more detail, using the mouse scroller. In the bottom left of the interface is the video player which provides play, pause and stop buttons, which enable the video to be precisely synchronized with the tri-axial accelerometer data streams. When creating an annotation/tag, users are required to input data including: the Creator, activityTag and Description in an annotation form displayed on the right hand side of the user interface. The successfully created annotations are stored on the RDF triple store and listed in the Annotation List.

Fig. 4 shows how users can retrieve specific annotations to train a SVM (C-SVC) activity classifier. It illustrates how a user searches and retrieves all annotations involving a human actor (with ID = “Jackie”) to train a low-level classifier.

Fig. 5 shows a screenshot of the results of applying a low-level human SVM activity classifier. This classifier identified six human activities (walking, running, sit-ups, lying, standing and sitting) and the result is shown in the plot visualization with the activity type tags displayed in blue along the top. The pie chart on the right shows the statistical information about each activity. From the pie chart, we can affirm that the participant spent 19.5% (132 s) of his time running, 18.9% (128 s) walking, 11.1% (75 s) doing sit-ups.

4. Evaluation and results

This section describes the evaluation methods that we employed to assess our system. Firstly we evaluated the performance of SAAR based on the results of our experiments on the human, dog and Eurasian badger data sets. Secondly we evaluated the usability of our system based on feedback from a group of 8 biologists.

4.1. Performance evaluation and classification results

The first experiment was conducted on human data with the aim of automatically tagging/identifying both high level (active and inactive) and low level activities (walking, running, sit-ups, standing, sitting and lying). Two classification models (a high level classifier and a low level classifier) were developed by feeding the training set (8 people, 16 min of data each) into the SVM (C-SVC algorithm). We then submitted the random human datasets (8 people, 15 min of data each) into the classifiers and compared the automatically generated results with the reference data (which was manually tagged using the video as reference).

The second experiment involved training a high level dog activity classifier to distinguish active and inactive dog behaviors and a low level dog activity classifier to identify walking, running, standing, sitting and lying movements on the dog data set. We then submitted the untagged datasets for dogs into the classifiers and compared the results with the reference data (manually tagged using video as reference).

The third experiment involved training a high level Eurasian badger activity classifier to identify high level activities (active and inactive), as well as a low level Eurasian badger activity classifier to recognize walking, running, climbing, foraging, standing and lying. We then submitted untagged datasets for Eurasian badgers into the classifiers and compared the results with the manually tagged reference data.

The last experiment involved using the high and low level classifiers generated from dog training data to automatically tag the corresponding Eurasian badger data. We then compared these results with the results from the third experiment to see if the dog classifier could successfully be used to recognize badger activities.

In previous studies, several methods have been proposed for assessing the performance of the supervised Machine Learning approach (Sokolova et al., 2006). In our study case, we used four commonly-accepted performance evaluation metrics which are calculated from the number of correctly and incorrectly recognized tags for each class. These metrics include true positive (TP), false positive (FP), true negative (TN) and false negative (FN). From these four metrics we calculate: accuracy ((TP+TN)/(TP+TN+FP+FN)), sensitivity (TP/(TP+FN)), precision (TP/(TP+FP)), and specificity (TN/(FP+TN)).

The performance results of the eight activity classification models are presented in Fig. 6. The results from the first three experiments (1), (2) and (3), reveal that the high level classification models produce: accuracy > 97%, sensitivity > 96%, precision > 97% and specificity > 96%. On the whole, these results are excellent and also...
better than the low level classifiers which produce: accuracy > 96%, sensitivity > 80%, precision > 80% and specification > 95%.

In addition, the human classification models performed better than the dog and badger classification models. The reason for this is that compared with dog and Eurasian badger data sets, the human data set contains less noise. A human-being is able to accurately perform the requested movements for the specified time period and the change from one activity to another is quite distinct. However, the animals (dogs and Eurasian badgers) are unable to perform specific movements/postures without professional training. Domestic dogs that were led by their owners were able to perform the requested range of activities much better than undomesticated badgers being monitored in the wild. Hence, the dog data contained less noise than the badger data and the dog classification engine (experiment 2) performed better than the badger classifier (experiment 3).

The results from the fourth experiment show that using the domestic dog classification model to recognize Eurasian badgers’ activities does not perform as well as the other three experiments—especially if we compare the high level (active/inactive) classifiers for experiment 4 against experiment 3 (whose classification model was generated from badger data). This is to be expected. However, the results are still quite positive. The high level classifier produced: accuracy > 92%, sensitivity > 85%, precision > 85% and specification > 87%, while the low level classifier produced: accuracy > 83%, sensitivity > 83%, precision > 79% and specification > 85%. To conclude, migrating the classification models across species does not perform as well as species-specific classification models, however in situations where there is no video reference, it can be used as an effective first pass, that can be corrected or refined manually by experts.

The other problem with migrating a classification model across species is that the activity terms/ontology may differ. For example, the dog activity ontology does not include the terms “foraging/digging” and “climbing” which are in the badger activity ontology. We specifically captured accelerometry data from the King Charles Spaniel while it was performing “foraging/digging” and “climbing” activities, because we knew in advance that we wanted to build a classifier that could be applied to the badger data. However, in general, there will not be a one-to-one mapping between terms in activity ontologies across species.

4.2. Usability evaluation

The usability of the system was assessed by users via both questionnaires and by observing users’ behavior during the usability test phase—which involved a small group of eight ecologists from the University of Queensland EcoLab. Users were asked to respond to the following questions on a questionnaire:

- I think the visualization interface is a useful tool.
- I found the visualization interface easy to use.
- I think the annotation interface is a useful tool.
- I found the annotation interface easy to use.
- I found the suggested tags appropriate.
- I found the search options useful.
- I found the search interface easy to use.
- I found it easy to train the automatic activity classification (SVM) engine.
- I felt confident using SAAR.
- I think my colleagues would learn SAAR quickly.
- I needed a lot of training before I could use SAAR effectively.
- The pie chart showing statistical information about each activity is useful.
Fig. 6. The experimental results from applying activity identification models to accelerometer datasets for different species.
Users were asked to respond to each question from a 5-point Likert scale ranging from “Strongly disagree” to “Strongly agree”. The system’s efficiency was determined by measuring the average time it took a user to: create an annotation through the timeline visualization, create an annotation through the video pane, search existing annotations to train a SVM classifier, and use the dynamically created classifier to predict animal activities. Each user was given a brief tutorial in the use of the SAAR system and then assigned a specific set of annotation and recognition tasks. The time taken to complete each task was recorded.

4.3. Usability evaluation results

The questionnaire results were very positive. All of the users who were surveyed found the visualization interface, annotation interface, search interface and the pie chart to be useful, and believe that their colleagues will learn to use the SAAR system quickly. 87.5% of users found that the system, including the visualization, annotation and search interface, is easy to use. 12.5% of users felt that they would require more time to learn to use it effectively. Aspects that required further information or clarification included, instructions on how to operate the zoom in and zoom out functions for the timeline visualization and explanations of the meaning of each search option.

Table 1 shows the results of the time trials. The average time taken by users to create an annotation through the timeline visualization was 19.5 s which was less than the time required to create an annotation using the video pane. This is because users typically had to re-play video segments multiple times (rewind, pause and replay) to be clear on what the animal was doing. The time to complete Tasks 3 and 4 both depends on the size of the data set but the average time for each task was very low (13 s and 6 s respectively). These times are considerably less than the time that would be required if animal behavior researchers had to implement their own SVM.

It was not possible to compare these times with other comparable systems because, to our knowledge, there are no other systems that support similar functionalities. However, we believe that these results indicate that users are able to use our system quickly and effectively to complete the required tasks and that in general, our system will significantly expedite the process of analyzing large volumes of 3D accelerometry data.

5. Limitations and future work

Despite the convincing results presented above, the user tests also revealed a number of system limitations. In particular, the system currently only supports accelerometer data of sampling rate 1 Hz. With some species, this sampling rate is insufficient to recognize specific activities. However, the zoom in and zoom out functionalities associated with the Plot panel became very slow when displaying higher sampling rates or large data sets e.g., 3–5 hour data sets. This is due to the fact that the visualization software redraws all the points when users zoom in or zoom out the Plot visualization. We need to investigate more scalable approaches for redrawing data files with sampling frequency > 1 Hz or data files for extended periods (> 2 h).

The second problem that we encountered was related to the quality of the data acquired from the accelerometers. It was quickly apparent that the integrity of the data is compromised if the position of the accelerometer changes during the data capture phase. The accelerometers must be rigidly attached in a fixed position and orientation on the back of the animal being monitored. If the position or the orientation changes, the quality of both the data and the classification deteriorates. This is one of the major challenges associated with accelerometers—especially when dealing with wild animals such as crocodiles or bears, it is difficult to always attach the accelerometer to precisely the same location and to ensure it does not shift over time.

We have also identified a number of future work directions that we would like to pursue. Firstly, from the system point view, we plan to integrate the Plot and Video visualization interfaces with the Google Map interface to enable simultaneous visualization of tri-axial accelerometer data streams, videos and GPS location information. In addition, we are also planning to develop an energy expenditure distribution map by analyzing both animal day and night movements to predict animal health statuses.

Secondly, a number of researchers are attaching GPS acoustic and satellite tags (Guo et al., 2009; Wagenaar et al., 2011) that track GPS location on a larger scale as well as other sensors (that measure body temperature, heart rate, bioacoustics, etc.) to animals, in addition to accelerometers. The SAAR system could be extended to support the integration, visualization and analysis of these additional parameters—in order to detect more complex behaviors, including interactions between animals, such as mating, fighting or territorial marking.

We are also hoping to acquire tri-axial accelerometer data captured from wild dingoes in the future. Dingoes are closer in species and behavior to domestic dogs, so we expect that the domestic dog classifier will work better on dingo data than it did on the badger data, but we would like to test this hypothesis. We are also interested in applying and evaluating this research to accelerometer data captured from flying foxes and birds to monitor their behavior.

To date, we have only evaluated SVMs using the C-SVC algorithm. It would be interesting to evaluate the classification results using different types of SVMs (e.g., nu-SVC, regressing SVM) and different kernel functions—to determine which SVM and kernel function produce the best results.

Currently the uploaded data sets are openly available via the SAAR Web site. However, many researchers would prefer to limit access to their experimental data only to project partners, at least until the data has been published. It also makes sense to restrict who can attach/edit tags on accelerometer data streams. In the future we plan to implement authentication and access control protocols over the datasets and associated tags, to enable the data owner to specify access controls including read, edit and re-use permissions.

Finally, once we have implemented the user authentication and access control protocols, we plan on making the system available to a broader community of users, beyond just the University of Queensland ECO-Lab. This will enable us to test the performance, scalability and usability of the system based on feedback from a larger volume of users.

6. Conclusion

We believe that the SAAR system delivers an easy-to-use Web-based repository and a set of semantic tagging, visualization and activity recognition services that will greatly benefit those researchers who are using accelerometers to quantify animal movement and behavior. The data visualization and semantic annotation/tagging interface enable rapid exploration and interpretation of the accelerometer datasets. More importantly the user interface enables domain experts to record their knowledge in a format that can easily be re-used to develop accurate machine learning algorithms capable of automatically recognizing behavioral patterns in a wide range of species.

Table 1

<table>
<thead>
<tr>
<th>Task description</th>
<th>Time range</th>
<th>Average time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Create a new annotation using the Plot timeline</td>
<td>15–32 s</td>
<td>19.5 s</td>
</tr>
<tr>
<td>2. Create a new annotation using the video pane</td>
<td>15 s–3.5 min</td>
<td>45.9 s</td>
</tr>
<tr>
<td>3. Search and retrieve annotations and input as training data to generate new classifier</td>
<td>3 s–1 min</td>
<td>13 s</td>
</tr>
<tr>
<td>4. Submit new 15 min dataset into classifier, generate automatic tags and display in visualization pane</td>
<td>4 s–1.5 min</td>
<td>6 s</td>
</tr>
</tbody>
</table>
Acknowledgments

The work presented in this paper is supported by the China Scholarship Council. The authors would also like to thank Professor Craig Franklin and the University of Queensland ECO-Lab for their valuable support, contributions and feedback. All procedures were carried out with the approval of the University of Queensland Animal Ethics Committee (SBS/300/12/NECTAR). Owen R. Bidder is part-funded by the European Social Fund (ESF) through the European Union’s Convergence program administered by the Welsh Government and by the Royal Society for the Prevention of Cruelty to Animals.

References